EXHIBIT 3

International Journal of Risk & Safety in Medicine 24 (2012) 221–231
DOI 10.3233/JRS-2012-0571
IOS Press

The paroxetine 352 bipolar trial: A study in medical ghostwriting

Jay D. Amsterdam[a,*] and Leemon B. McHenry[b]
[a]*Depression Research Unit, Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA, USA*
[b]*Department of Philosophy, California State University, Northridge, CA, USA*

Abstract. *Background:* The problem of ghostwriting in corporate-sponsored clinical trials is of concern to medicine, bioethics, and government agencies. We present a study of the ghostwritten archival report of an industry-sponsored trial comparing antidepressant treatments for bipolar depression: GlaxoSmithKline (GSK) paroxetine study 352. This analysis is based upon publicly available evidence presented in a complaint of research misconduct filed with the Office of Research Integrity of the Department of Health and Human Services.

Objectives: We performed a deconstruction of the published study to show how primary and secondary outcome analyses were conflated, turning a 'negative' clinical trial into a 'positive' study – with conclusions and recommendations that could adversely affect patient health.

Methods: The paroxetine 352 study was a randomized, double-blind, placebo-controlled, 19-site trial comparing paroxetine and imipramine in 117 patients with bipolar type I major depressive episode which was unresponsive to prior lithium carbonate therapy.

Results: Analysis of the primary outcome measures found no statistically significant difference between paroxetine or imipramine versus placebo. However, the published article concluded that both drugs were efficacious versus placebo for a *post hoc* subgroup of patients.

Conclusions: Few industry-sponsored studies gain public scrutiny. It is important to make these articles transparent to the scientific and medical community.

Keywords: Ghostwriting, key opinion leaders, depression, bipolar, SSRI, paroxetine, litigation, industry sponsorship

1. Introduction

The problem of truth and transparency in published scientific reports of corporate-sponsored clinical trials has been an on-going concern in the medical and bioethics literature. The difference between what a trial should report and what is actually reported in the medical journals in the past 30 years is so alarming that some editors have declared a crisis of credibility [1]. Details of selective data reporting, misrepresentation of results, and ghostwriting of manuscripts have been revealed from the fragmentary release of documents publicly disclosed in litigation in the United States [2]. Critical evaluation of these practices, however, is sporadic at best because of inherent non-disclosure or inaccessibility of

*Address for correspondence: Jay D. Amsterdam, MD, Depression Research Unit, University Science Center – 3rd Floor, 3535 Market Street, Philadelphia, PA 19104, USA. Tel.: +1 215 662 3462; Fax: +1 215 662 6443; E-mails: enopath@aol.com, jamsterd@mail.med.upenn.edu.

information that remains under judicial seal by the courts. Even the most egregious cases of corporate fraud in the pharmaceutical industry are protected from public disclosure and only become available when attorneys challenge the confidentiality status of the documents. Another, but less common, source of information about unethical reporting and publishing comes from whistleblower complaints where honest investigators refuse to compromise scientific accuracy for commercial or institutional objectives [3].

This report describes how a 'negative' clinical trial was published as a 'positive' study with unsubstantiated claims of efficacy and safety. It is based upon public evidence presented in a complaint of research misconduct filed with the Office of Research Integrity (ORI) of the Department of Health and Human Services (HHS) [4]. Additional supporting documents are available, but remain under current court seal. The article, entitled "Double-Blind, Placebo-Controlled Comparison of Imipramine and Paroxetine in the Treatment of Bipolar Depression," was published in the *American Journal of Psychiatry* in June 2001 [5]. It appears to be part of a marketing effort by SmithKline Beecham (SKB), now GlaxoSmithKline (GSK), to have paroxetine (Paxil®/Seroxat®) outsell all of the competition in the selective serotonin reuptake inhibitor (SSRI) antidepressant market for all indications [6]. This strategic plan included the ghostwriting of articles by medical communication companies that would 'spin' the data in favor of the sponsor's product. For several of these unapproved, but much-sought after indications, we believe that ghostwritten articles were placed in leading medical journals to facilitate off-label prescriptions by clinicians [7]. A key component of this *sub rosa* plan was for prominent academic researchers, known in industry as 'key opinion leaders,' to lend their names to these publications as authors, which would give the appearance of scientific objectivity.

It is difficult to believe that the publication of the paroxetine 352 study was for any other purpose than to facilitate the prescription of paroxetine for the treatment of the depressive phase of manic depressive, or bipolar, disorder (i.e., bipolar depression). This area of treatment represented a natural extension of the already approved indication for paroxetine of unipolar major depressive disorder (or non-bipolar depression). However, despite the potentially lucrative aspect of this indication for marketing paroxetine, it was also well known in the psychiatry community that antidepressants (including SSRIs) could worsen bipolar disorder by precipitating manic episodes [8–10] and increasing the risk of suicide in patients [11–13].

In this report, we deconstruct the paroxetine 352 study and show how the ghostwriters conflated the primary and secondary outcome analyses that turned a 'negative' study into a 'positive' endorsement of antidepressant therapy (in particular paroxetine) for patients with bipolar depression.

2. Methods

2.1. Initial study design

Deconstruction of the paroxetine 352 study is based upon information obtained from the GSK Clinical Trials Website Result Summary for Study 29060/352 updated 09 March 2005 [14], GSK Paroxetine Protocol PAR-29060/352 (amended 22 July, 1994), Drug Industry Document Archives Website [15], and evidence presented in a complaint of research misconduct filed with the federal ORI filed on July 8, 2011 by Dr. Jay D. Amsterdam, Professor of Psychiatry at the University of Pennsylvania [4]. The paroxetine 352 study was conducted between February 1994 and March 1996. It was originally designed as a 10-week, 18-site, randomized, double-blind, placebo-controlled comparison of paroxetine versus imipramine in subjects with bipolar type I disorder and was designated a Phase IV (i.e., post-marketing,

non-indication) study with a projected duration of 2 years. Its objective was "to compare the efficacy and safety of paroxetine and imipramine to [placebo] in the treatment of bipolar depression in subjects stabilized on lithium therapy [14]". The primary efficacy measures, as described in the study protocol, were the change from baseline in the 17-item Hamilton Rating Scale for Depression (HAM-D) total score [16], and the change from baseline in the Clinical Global Impression Severity of Illness (CGI-S) score [17] for paroxetine versus placebo and for imipramine versus placebo. The stated secondary efficacy measures were the proportion of subjects with a final HAM-D score ≤ 7 or a final CGI-S score ≤ 2. Additional secondary outcomes included the proportion of subjects experiencing adverse events, premature treatment discontinuation, and manic or hypomanic reactions as determined by the DSM-III-R Mania/Hypomania Assessment and the Young Mania Rating Scale (YMRS) [18]. A post-treatment safety assessment was performed after subjects were discontinued from double-blind medication to determine the proportion of subjects in each treatment group with discontinuation symptoms due to adverse events (e.g., withdrawal and/or manic symptoms).

The study population consisted of outpatient subjects ≥ 18 years old, with a lifetime diagnoses of bipolar type I disorder and a history of at least one prior manic or major depressive episode within the preceding 5 years. Subjects had a current major depressive episode who failed to respond to lithium carbonate therapy (or lithium carbonate plus sodium valproate or carbamazepine therapy) for ≥ 7 weeks with a serum lithium level of 0.5–1.2 mEq/L for ≥ 6 weeks prior to starting the study. Treatment procedures for the study have been published [5, 14].

2.2. Sample size estimate

The original protocol called for a sample size of 62 subjects per treatment group (or a total of 186 subjects). However, during the course of the study the protocol sample size estimate was formally amended downward to 46 subjects per treatment group (or a total of 138 subjects) [14].

2.3. Statistical methods

A summary of the statistical plan in the amended study protocol called for analyses to be performed on two sets of efficacy data (i.e., interim visit-wise data and endpoint data), with endpoint data considered the primary data set. Unspecified statistical procedures were to be applied to the outcome data to remove potential rating bias that might occur in subjects with concurrent depressive and manic symptoms. Separate analyses were to be performed on the entire subject population, and on two subgroups of subjects: (i) those who experienced a manic or hypomanic episode during the study; and, (ii) those who did not experience a manic or hypomanic episode during the study. The YMRS was to be used to assess severity of manic and/or hypomanic symptoms across treatment groups, and the relationship between change from baseline in YMRS scores and HAM-D scores was to be specifically examined.

Factors that might influence treatment outcome were to be examined via the use of interaction terms in the regression models (i.e., treatment, investigator site, strata of "high" or "low" baseline lithium level). However, it was specifically noted that any interaction term that was not statistically significant (i.e., $p > 0.1$) in the primary analysis would be dropped from all subsequent analyses. The protocol noted that the comparison of primary interest is paroxetine versus placebo across (regardless of) lithium strata, performed at a two-tailed significance level of $p = 0.05$. The protocol also required that differences among subject groups in baseline demographic and diagnostic variables would be checked and, if differences existed for variables predictive of response, their impact on the results would be investigated.

Finally, mania and hypomania were to be analyzed using logistic regression models that included the effect terms of 'treatment', 'investigator', and 'treatment × investigator' interaction. The protocol noted that if the interaction was not significant, it would be dropped from the model.

3. Results

3.1. Changes in original study methodology and reporting

The original protocol sample size estimate of 0.9 (1-β) or 62 subjects per treatment group was officially amended downward to 0.8 (1-β) or 46 subjects per group during the study. The latter value was the sample size described in the GSK Clinical Trials Website Result Summary [14]. No explanation was provided for this change in sample size in the amended protocol. However, we suspect that this reduction in power might have resulted from exceedingly slow subject recruitment into the study, which ultimately led GSK to add a 19th investigative site. By the time GSK decided to halt subject enrollment prematurely and terminate the study, only 117 (of the originally projected 186 subjects) were enrolled, resulting in final sample sizes for paroxetine ($n = 35$), imipramine ($n = 39$), and placebo ($n = 43$). By the time the study was published in June 2001 in the *American Journal of Psychiatry*, however, the declared sample size estimate had again changed with the article stating: "The study was designed (*sic*) to enroll 35 patients per arm, which would allow 70% power to detect a 5-point difference on the Hamilton depression scale score (SD = 8.5) between treatment groups [5]".

Although the published article noted that statistical power was estimated at only 70%, the article did not inform the reader that this value represented an unconventionally low power for a clinical trial. The article did not inform the reader that the original power estimate was 62 subjects per group or that the original power estimate had been officially reduced during the study. Moreover, the article made no mention of the fact that the final power estimate was determined after the study was completed, and that this *post hoc* power estimate most likely occurred as an 'extra-regulatory' protocol change in order to allow the final sample size estimate of 35 subjects per group to comport with the final sample size of the paroxetine group (i.e., $n = 35$). The published article failed to acknowledge clearly that the study failed to recruit the projected sample size necessary to test the primary study hypothesis, and only hinted by its published sample size estimate that the study had insufficient statistical power to test the primary study aims.

The statistical plan described in the GSK Clinical Trials Website Results Summary [14] was considerably less detailed and somewhat different than that of the original or amended study protocol. It presented the primary comparison of interest as paroxetine versus placebo (regardless of baseline lithium level strata). It noted that comparisons of secondary interest were imipramine versus placebo and imipramine versus paroxetine. The statistical plan concluded with the statement: "In addition, analyses were performed within lithium strata, which included only an effect for treatment". The GSK Clinical Trials Website provided no indication that these secondary analyses were performed as *post hoc* procedures, or that they were not statistically warranted – as the primary outcome analyses showed no significant treatment × lithium level strata interaction effect for any of the treatment groups [14].

The analysis plan of the published article played down the statistical procedures used for the primary efficacy analyses and, instead, emphasized the procedures used for analyzing the *post hoc* lithium level strata efficacy analyses. No mention was made in the published article indicating that the sample size of 35 subjects per group represented insufficient statistical power to test adequately for differences among

lithium level subgroups [5]. Moreover, the statistical plan presented in the published article noted that "no adjustments for multiple comparisons were made [5]" which, if properly applied, would have nullified the statistical significance of the only two 'positive' comparisons found in the study.

Finally, in contrast to the amended study protocol that called for the analysis of safety measures to examine manic symptoms (i.e., DSM-III-R Mania/Hypomania Assessment and YMRS), no mention of these measures was made in the GSK Clinical Trials Web-site Results Summary or in the published article, and no safety data acquired from these measures were presented in either published venue.

3.2. Conflation of primary and secondary outcome measures

We believe that GSK conflated primary and secondary study aims, and then presented 'positive' *post hoc* lithium level strata analyses as if they were the primary analyses of interest (Table 1) [14]. Of the more than 30 separate efficacy analyses reported in the GSK Clinical Trials Web-site Results Summary [14], only two *post hoc* comparisons showed statistical significance (i.e., paroxetine versus placebo for change in HAM-D score in the "low" lithium level group ($p = 0.049$) and imipramine versus placebo for change in HAM-D score in the "low" lithium level group ($p = 0.038$) (Table 1).

The published article provided a cursory statement indicating that the primary outcome analysis was 'negative'. However, we believe that the article then conflated the primary and secondary analyses by attributing the 'negative' primary finding to an excessive placebo response rate in the "high" lithium level subgroup (although there is no discernible evidence in the study data to support this conclusion). After conflating the primary and secondary outcome measures, the published article then emphasized the only 'positive' efficacy finding for paroxetine as if it was the primary study aim: "... among the low serum lithium level patients, paroxetine and imipramine were superior to placebo in terms of mean change from baseline in scores on the Hamilton depression scale and CGI severity of illness scale... [5]".

In addition, we believe that the published article conflated efficacy and adverse event data to favor paroxetine over imipramine. For example, the article presented only selected safety data on treatment-emergent manic episodes that favored paroxetine. In contrast to the study protocol that called for the analysis of manic and hypomanic symptoms as measured by the DSM-III-R Mania/Hypomania Assessment and YMRS, no data from these safety measures were presented in either the GSK Clinical Trials Website Results Summary or in the published article. Instead, the GSK Clinical Trials Website Results Summary and the published article simply noted the number of manic episodes that were clinician-reported during the trial and appeared to emphasize the lack of manic episodes reported with paroxetine (versus imipramine and placebo).

We believe that the published article also favored the adverse event profile of paroxetine by portraying it as having virtually no sexual side effects (versus imipramine) with the article stating: "Patients treated with imipramine reported a higher incidence of abnormal ejaculation (18.8%) and impotence (25.0%) than did patients receiving paroxetine (0.0 and 6.3%, respectively) or placebo (5.0 and 0.0%, respectively) [5]". By conflating the 'positive' efficacy and favorable side effect profile of paroxetine, the article suggested that paroxetine was the obvious drug of choice for physicians to prescribe – especially when the article emphasized the greater sexual side effect burden and manic switch rate with imipramine and placebo.

However, closer examination of the available data suggests that this conclusion may be unwarranted. For example, although the published article noted that the most common side effects occurring with paroxetine (at a frequency $\geq 10\%$) were insomnia and somnolence, the published article failed to report a higher frequency of paroxetine-induced treatment-emergent depression (versus imipramine) – as listed in the GSK Clinical Trials Website Results Summary (Table 2). The higher frequency of insomnia and

Table 1
Primary efficacy results derived from the GSK Clinical Trials Website Result Summary [14]

Primary efficacy results (Intent to treat analysis): Baseline mean and change from baseline mean for the first 17-items of HAMD-21 LOCF	Paroxetine (Par)	Imipramine (Imp)	PBO
Total response dataset	*N* = 33	*N* = 36	*N* = 43
Baseline, *n* mean HAMD-21 (se)	33 20.38 (0.68)	36 20.71 (0.65)	43 21.57 (0.59)
Baseline pairwise comparison treatment *p*-values	Par vs. PBO 0.191	Imp vs. PBO 0.332	Par vs. Imp 0.725
Endpoint, *n* change from baseline mean (se)	33 –10.2 (1.27)	36 –10.1 (1.21)	43 –8.06 (1.11)
Endpoint pairwise comparison treatment *p*-values	Par vs. PBO 0.199	Imp vs. PBO 0.220	Par vs. Imp 0.932
High lithium level dataset	*N* = 14	*N* = 17	*N* = 21
Baseline, *n* mean (se)	14 20.29 (1.01)	17 21.35 (0.91)	21 21.95 (0.82)
Baseline pairwise comparison treatment *p*-values	Par vs. PBO 0.206	Imp vs. PBO 0.628	Par vs. Imp 0.436
Endpoint, *n* change from baseline mean (se)	14 –9.79 (1.90)	17 –9.35 (1.72)	21 –10.4 (1.55)
Endpoint pairwise comparison treatment *p*-values	Par vs. PBO 0.809	Imp vs. PBO 0.659	Par vs. Imp 0.867
Low lithium level dataset	*N* = 19	*N* = 19	*N* = 22
Baseline, *n* mean (se)	19 20.37 (0.92)	19 20.11 (0.92)	22 21.18 (0.85)
Baseline pairwise comparison treatment *p*-values	Par vs. PBO 0.519	Imp vs. PBO 0.394	Par vs. Imp 0.840
Endpoint, *n* change from baseline mean (se)	19 –10.4 (1.67)	19 –10.7 (1.67)	22 –5.82 (1.56)
Endpoint pairwise comparison treatment *p*-values	Par vs. PBO 0.049	Imp vs. PBO 0.038	Par vs. Imp 0.912
Baseline mean and change from baseline mean for CGI-S LOCF	Paroxetine	Imipramine	PBO
Total response dataset	*N* = 33	*N* = 36	*N* = 43
Baseline, *n* mean (se)	33 4.21 (0.12)	36 4.31 (0.11)	43 4.33 (0.10)
Baseline pairwise comparison treatment *p*-values	Par vs. PBO 0.476	Imp vs. PBO 0.897	Par vs. Imp 0.573
Endpoint, *n* mean (se)	33 –1.33 (0.24)	36 –1.28 (0.23)	43 –0.91 (0.21)
Endpoint Pairwise comparison treatment *p*-values	Par vs. PBO 0.196	Imp vs. PBO 0.245	Par vs. Imp 0.876
High lithium level dataset	*N* = 14	*N* = 17	*N* = 21
Baseline, *n* mean (se)	14 4.21 (0.16)	17 4.35 (0.15)	21 4.29 (0.13)
Baseline pairwise comparison treatment *p*-values	Par vs. PBO 0.738	Imp vs. PBO 0.739	Par vs. Imp 0.535
Endpoint, *n* change from baseline mean (se)	14 –1.14 (0.38)	17 –0.94 (0.35)	21 –1.24 (0.31)
Endpoint pairwise comparison treatment *p*-values	Par vs. PBO 0.848	Imp vs. PBO 0.528	Par vs. Imp 0.698
Low lithium level dataset	*N* = 19	*N* = 19	*N* = 22
Baseline, *n* mean (se)	19 4.21 (0.17)	19 4.26 (0.17)	22 4.36 (0.16)
Baseline pairwise comparison treatment *p*-values	Par vs. PBO 0.516	Imp vs. PBO 0.670	Par vs. Imp 0.829
Endpoint, *n*	19	19	22
Change from baseline mean (se)	–1.47 (0.31)	–1.58 (0.31)	–0.59 (0.29)
Endpoint pairwise comparison treatment *p*-values	Par vs. PBO 0.040	Imp vs. PBO 0.022	Par vs. Imp 0.810

depression in the paroxetine-treated subjects could have represented drug-induced manic, hypomanic, or mixed manic and depressive symptoms. However, the notable lack of data analyses presented on manic and hypomanic symptoms from the DSM-III-R Mania/Hypomania Assessment and the YMRS in the published article make this possibility difficult to resolve. Moreover, the complete absence of manic or hypomanic symptoms reported with paroxetine stands in stark contrast to the vast majority of published research articles on antidepressant-induced mood conversion symptoms in bipolar depression.

Table 2
Number (%) of reported and elicited adverse events derived from the GSK Clinical Trials Website Result Summary [14]

Most frequent adverse events	Paroxetine ($n=35$)	Imipramine ($n=39$)	Placebo ($n=43$)
Any adverse event	33 (94.3)	36 (92.3)	40 (93.0)
Tremor	14 (40.0)	15 (38.5)	4 (9.3)
Insomnia	13 (37.1)	5 (12.8)	10 (23.3)
Somnolence	12 (34.3)	13 (33.3)	11 (25.6)
Diarrhea	10 (28.6)	6 (15.4)	7 (16.3)
Nausea	9 (25.7)	12 (30.8)	5 (11.6)
Headache	9 (25.7)	16 (41.0)	17 (39.5)
Infection	6 (17.1)	2 (5.1)	1 (2.3)
Dizziness	5 (14.3)	11 (28.2)	1 (2.3)
Constipation	5 (14.3)	11 (28.2)	4 (9.3)
Asthenia	5 (14.3)	5 (12.8)	2 (4.7)
Dry Mouth	4 (11.4)	24 (61.5)	3 (7.0)
Depression	4 (11.4)	0	4 (9.3)
Dyspepsia	2 (5.7)	7 (17.9)	1 (2.3)
Sinusitis	2 (5.7)	2 (5.1)	4 (9.3)
Myalgia	1 (2.9)	0	5 (11.6)
Vomiting	0	6 (15.4)	1 (2.3)
Back pain	0	2 (5.1)	5 (11.6)
Serious adverse events*	0	2 (5.1)	4 (9.3)
Mania	0	1 (2.6)	0
Aggression/homicidal ideation	0	1 (2.6)	0
Manic episode	0	0	2 (4.7)
Paranoia, hallucinations, delusions	0	0	1 (2.3)
Re-emergence of depression	0	0	1 (2.3)
Subjects with fatal SAEs, *n* (%)	0	0	0

*Includes both fatal and non-fatal events.

4. Discussion

4.1. Misrepresentation and bias in reporting study outcomes

The recruitment of only 117 (of the originally projected 186) subjects resulted in a failed (i.e., non-informative) trial with inconclusive results. As a consequence, we believe that the published article had to rely on conflated *post hoc* analyses of data subsets to portray a favorable result for paroxetine. This favorable result was identified by subdividing the already insufficient sample size of each treatment group into smaller subgroups of "high" (i.e., 0.8–1.2 mEq/L) versus "low" (i.e., 0.4–0.79 mEq/L) lithium level strata. This procedure resulted in 6 treatment subgroups: three "high" lithium level subgroups of paroxetine ($n=14$), imipramine ($n=17$), and placebo ($n=21$) and three "low" lithium level subgroups of paroxetine ($n=19$), imipramine ($n=19$), and placebo ($n=22$). A pairwise comparison of treatments within these limited subgroups then identified two 'positive' findings: paroxetine superior to placebo in the "low" lithium level subgroup ($p=0.049$) and imipramine superior to placebo in the "low" lithium level subgroup ($p=0.038$). We believe that these 'positive' findings were then presented as if they were the

main finding of the study. A conflation of the 'positive' efficacy finding for paroxetine with the favorable side effect profile presented for paroxetine provided a positive 'spin' for paroxetine and a negative 'spin' for imipramine in the published article.

The published article failed to disclose that the statistical power of the study was insufficient to determine whether or not paroxetine (or imipramine) was actually superior to placebo in the "low" lithium level subgroup. The article failed to acknowledge that the study was not designed to test whether or not paroxetine (or imipramine) was superior to placebo in subjects with "high" or "low" lithium levels. To test this hypothesis, study subjects would have needed to be maintained in separate lithium level cohorts of "high" or "low" lithium level ranges throughout the study. In fact, the study merely stratified subjects into "high" or "low" lithium level ranges based upon a single baseline lithium level determination for statistical purposes. There were no discrete "high" or "low" lithium levels subgroups in the 352 study, as all study subjects were maintained (according to protocol) within a lithium level range of 0.4–1.2 mEq/L. We believe that it was the conflation of primary and secondary analyses that allowed GSK to present the lithium subgroup analyses as if they were distinct and clinically meaningful entities.

The article failed to inform the reader that the published power estimate of 35 subjects per group was not part of the original study design. Rather, in our opinion, the published power estimate was most likely contrived by GSK to comport with the final sample size of the paroxetine treatment group (i.e., 35 subjects). All of the academic authors named on the published article were also listed as clinical investigators on the trial. Thus, they should have been aware of the unconventional power estimate used in the published article, and that it differed from that of the study protocol. Authors need to exercise due diligence when reviewing a draft manuscript. Had they the opportunity to do so in this case, they should have been able to recognize that the final power estimate was substantially lower than that of the study protocol. Moreover, they would have recognized that the power estimate changed after the trial was completed. This suggests to us that the final change in power estimate was made by GSK without regulatory approval or oversight. If so, this change would likely represent a substantial departure from Good Clinical Practice Guidelines policies for the conduct of clinical trials in humans [19].

By setting forth the fact that there was a clinically meaningful distinction between "high" and "low" lithium level subgroups, the published article provided the reader with a false impression that patients with "low" lithium levels might be uniquely responsive to paroxetine. In addition, the article also asserts a high rate of sexual side effects with imipramine, while reporting almost no sexual side effects with paroxetine. This claim, even if true in the 352 study, is misleading and implies that paroxetine is likely to cause few, if any, sexual side effects in bipolar patients. To enhance this proposition, the article selectively cited only side effect literature on imipramine, while omitting any citation on sexual side effects with paroxetine.

Finally, in our opinion the published article is misleading in its assertion that there is no therapeutic advantage to using antidepressant therapy in bipolar depressed patients with "high" lithium levels. The 352 study was not designed to test this hypothesis and this conclusion does not appear to be supported by the data. Moreover, with the exception of the paroxetine 352 study, there are no other published studies reporting a lack of antidepressant efficacy in patients with "high" lithium levels. Conversely, the assertion that antidepressants may be more effective in patients with "low" lithium levels is potentially dangerous and inconsistent with most published practice guidelines for treating bipolar depression [20, 21]. In this regard, patients with "low" lithium levels would be at greater risk for developing antidepressant-induced mania and suicidal ideation [8–13]. By downplaying the well-known side effect profile of paroxetine and portraying it as being effective for bipolar depression without manic episodes, the published article was able to conflate successfully efficacy and side effect aims to favor paroxetine over imipramine. It is our

opinion that the article's recommendation that "patients with bipolar depression who maintain high serum lithium levels may not require additional antidepressant medications... [while] patients with low serum lithium levels or those who cannot tolerate high serum lithium levels may benefit from augmentation therapy with either paroxetine or imipramine" is speculative, misleading and is not supported by the study results.

4.2. The ghostwriting of paroxetine study 352

The paroxetine 352 study manuscript was ghostwritten by Scientific Therapeutic Information, Inc. (STI) with funds provided by GSK; however, neither STI, the ghostwriter, Sally Laden, nor GSK's role in the production of the manuscript was acknowledged in the published article. Prominent academic researchers (with financial ties to GSK) and GSK employees were designated by GSK as 'authors' on the manuscript. According to available evidence that was used in support of the ORI Complaint and subsequently disclosed in a follow up letter to ORI [22], the manuscript's authors were chosen by GSK in consultation with STI. This practice, and specifically STI's role in creating ghostwritten manuscripts, was criticized in an editorial in the *Journal of the American Medical Association* in relation to Merck & Co., Inc.'s effort to promote sales of Rofecoxib (Vioxx®) [23].

Many of the named authors on the published article had little or no direct involvement in the design, daily conduct, data analysis, or writing of the initial manuscript drafts. In fact, some of the authors were only selected for this role once the ghostwriters began to draft the manuscript from the final study report or a summary provided by GSK. It appears from the available evidence that GSK and STI had originally chosen Dr. Laszlo Gyulai, then Assistant Professor at the University of Pennsylvania, as the paper's first author [22]. However, Dr. Gyulai was subsequently removed from this position by GSK and replaced by two other authors who were assigned by GSK to the first and second positions on the paper [22]. The evidence also indicates that the final GSK-assigned authors on the published article never reviewed or even saw preliminary drafts of the paper, and only saw the final edited manuscript just prior to final acceptance by the *American Journal of Psychiatry* [22]. Conduct of that sort does not comport with the International Committee of Medical Journal Editors' standards of genuine authorship [24]. As we see it, GSK, STI, the named authors on the published article and, indeed, the *American Journal of Psychiatry* are all morally culpable in various degrees for complicity in publishing this article that we believe to be drug promotion disguised as science.

The 352 study invites comparison with another GSK-sponsored study (i.e., the paroxetine 329 study) in which paroxetine was compared to imipramine and placebo in children and adolescents with major depression [25]. Other than the *ad hoc* assignment of authors in the paroxetine 352 study, we see little difference in the manner in which both study reports were ghostwritten and published in the same year. In each case, the article failed to report all the relevant data resulting in a 'spin' in favor of paroxetine. A more detailed description of the ghostwriting of study 329 is now possible via documents de-classified in litigation [26], whereas in the case of the 352 study, GSK and STI have refused to release the available documents into the public domain on the basis of the alleged necessity of protecting trade secrets.

5. Conclusion

Corporate malfeasance in misrepresenting the results of clinical trials, especially where ghostwriting is involved, is of particular concern in the field of psychiatry where outcome measures of clinical trials

are more subjective and lend themselves to manipulation. We are convinced by our analysis that the GSK-sponsored paroxetine 352 study is one such instance of this practice, although prior examples have also come to light [27–30]. Because few industry-sponsored studies gain public scrutiny and even fewer are ever formally retracted [31], it is important to make these articles as transparent as possible in order to correct the scientific record and inform the medical community of potential harm.

Acknowledgments

The authors would like to thank Dr. Bernard Carroll and an anonymous reviewer for the journal for helpful suggestions to an earlier draft of this paper, and to Ronald Goldman and Skip Murgatroyd for legal review. The views expressed herein are those of the authors alone and not necessarily those of any other person, firm or entity.

Conflict of interest

Dr. Amsterdam was a clinical investigator at the University of Pennsylvania for paroxetine study 352 that was supported by a grant from GSK from 1994 to 1996. Dr. Amsterdam is currently not a member of any industry-sponsored advisory board or speaker's bureau, and has no financial interest in any pharmaceutical or medical device company. He has received legal support for his ORI Complaint from the law firm of Baum, Hedlund, Aristei & Goldman of Los Angeles, California.

Dr. McHenry is a member of Healthy Skepticism and research consultant for Baum, Hedlund, Aristei & Goldman.

References

[1] Smith R. Medical journals are an extension of the marketing arm of pharmaceutical companies. PLoS Med. 2005;2(5):e138.
[2] Kesselheim AS, Avorn J. The role of litigation in defining drug risks. JAMA. 2007;297(3):308–31.
[3] Lenzer J. What can we learn from medical whistleblowers? PLoS Med. 2005;2(7):e209.
[4] Complaint of Scientific Misconduct against Dwight Evans, Laszlo Gyulai, Charles Nemeroff, Gary S. Sachs and Charles Bowden, http://psychrights.org/Research/Digest/Science4Sale/110708EthicsComplaintAgainstEvansGyualaiNemeroffSachsBowdenetal.pdf (Accessed May 2012).
[5] Nemeroff CB, Evans DL, Gyulai L, Sachs GS, Bowden CL, Gergel IP, Oakes R, Pitts CD. Double-blind, placebo-controlled comparison of imipramine and paroxetine in the treatment of bipolar depression. Am J Psychiatry. 2001;158:906–12.
[6] http://abcnews.go.com/images/Primetime/paxil_bpg.pdf (Accessed May 2012).
[7] Nemeroff CB, editor. Advancing the treatment of mood and anxiety disorders: The first 10 years' experience with paroxetine. Psychopharm Bul. 2003;37/1.
[8] Wehr TA, Sack DA, Rosenthal NE, Cowdry RW. Rapidcycling affective disorder: Contributing factors and treatment responses in 51 patients. Am J Psychiatry. 1988;145(2):179-84.
[9] Sachs GS, Lafer B, Stoll AL, Banov M, Tibault AB, Tohen M, Rosenbaum JF. A double-blind trial of bupropion versus desipramine for bipolar depression. J Clin Psychiatry. 1994;55:391-3.
[10] Leverich GS, Altshuler LL, Frye MA, Suppes T, McElroy SL, Keck PE Jr, Kupka RW, Denicoff KD, Nolen WA, Grunze H, Martinez MI, Post RM. Risk of switch in mood polarity to hypomania or mania in patients with bipolar depression during acute and continuation trials of venlafaxine, sertraline, and bupropion as adjuncts to mood stabilizers. Am J Psychiatry. 2006;163:232-9.
[11] Dunner DL. Sub-types of bipolar affective disorder with particular regard to bipolar II. Psychiat Developments. 1983;1:75-85.

[12] Goodwin FK, Jamison KR. Manic-Depressive Illness. Bipolar Disorders and Recurrent Depression. 2nd ed. New York: Oxford University Press; 2007.
[13] Coryell W, Scheftner W, Keller M, Endicott J, Maser J, Klerman GL. The enduring psychosocial consequences of mania and depression. Am J Psychiatry. 1993;150:720-7.
[14] http://www.gsk-clinicalstudyregister.com/result_detail.jsp?protocolId=29060%2F352&studyId=F1D83A94-2628-4C9C-83A5-11D00A2D30AC&compound=paroxetine) (Accessed April 2012).
[15] http://dida.library.ucsf.edu/ (Accessed May 2012).
[16] Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry. 1960;23:56-62.
[17] Guy W, editor. ECDEU Assessment Manual for Psychopharmacology: Publication ADM 76-338, Washington, DC: US Department of Health, Education, and Welfare; 1976. pp. 218-22.
[18] Young RC, Biggs JT, Ziegler VE, et al. A rating scale for mania: Reliability, validity, and sensitivity. Brit J Psychiatry. 1978;133:429-35.
[19] US Code of Federal Regulations, Consolidated Good Clinical Practice (GCP) Guidelines, US DHHS, Food & Drug Administration Center for Drug Evaluation and Research (CDER), Washington, DC; April 1, 2002.
[20] American Psychiatric Association: Practice guidelines for 1994;151(Suppl):1-36.
[21] American Psychiatric Association Practice guidelines for the treatment of patients with bipolar disorder (revision). Am J Psychiatry. 2002;159(Suppl):1-50.
[22] http://blogs.nature.com/news/files/2012/06/Jay-Amsterdam-today-filed-a-24-page-complaint-with-the-Office-of-Research-Integrity-at-the-US-National-Institutes-of-Health.pdf (Accessed August 2012).
[23] DeAngelis CD, Fontanarosa PB. Impugning the integrity of medical science: The adverse effects of industry influence. JAMA. 2008;299/15:1833-5.
[24] http://www.icmje.org. (Accessed June 2012).
[25] Keller MB, Ryan ND, Strober M, Klein RG, Kutcher SP, Birmaher B, et al. Efficacy of paroxetine in the treatment of adolescent major depression: A randomized, controlled trial. J Am Acad Child Adolesc Psychiatry. 2001;40:762-72.
[26] McHenry L, Jureidini J. Industry-sponsored ghostwriting in clinical trial reporting: A case study, Account Res. 2008;15:152-67.
[27] Steinman MA, Bero LA, Chren MM, Landefeld CS. Narrative review: The promotion of gabapentin: An analysis of internal industry documents. Ann Inter Med. 2006;145:284-93.
[28] Ross JS, Hill KP, Egilman DS, Krumholz HM. Guest authorship and ghostwriting in publications related to rofecoxib: A case study of industry documents from rofecoxib litigation. JAMA. 2008;299/15:1800-12.
[29] Jureidini J, McHenry L, Mansfield P. Clinical trials and drug promotion: selective reporting of study 329. Int J Risk Saf Med. 2008;20(1-2):73-81.
[30] Fugh-Berman AJ. The haunting of medical journals: How ghostwriting sold "HRT". PLoS Med. 2010;7(9):e1000335. doi:10.1371/journal.pmed.1000335.
[31] Newman M. The rules of retraction. BMJ. 2010;341:1246-8.

Amsterdam, JD and McHenry, LB. 2019. *Journal of Scientific Practice and Integrity*. 1(1). DOI: 10.35122/jospi.2019.958452

Research Article

The Paroxetine 352 Bipolar Study Revisited: Deconstruction of Corporate and Academic Misconduct

Jay D. Amsterdam[i] and Leemon B. McHenry[ii]

Abstract

Medical ghostwriting is the practice in which pharmaceutical companies engage an outside writer to draft a manuscript submitted for publication in the names of "honorary authors," typically academic key opinion leaders. Using newly-posted documents from paroxetine litigation, we show how the use of ghostwriters and key opinion leaders contributed to the publication of a medical journal article containing manipulated outcome data to favor the proprietary medication. The article was ghostwritten and managed by SmithKline Beecham, now GlaxoSmithKline (GSK) and Scientific Therapeutics Information, Inc. without acknowledging their contribution in the published article. The named authors with financial ties to GSK had little or no direct involvement in the paroxetine 352 bipolar trial results and most had not reviewed any of the manuscript drafts. The manuscript was originally rejected by peer review; however, its ultimate acceptance to the *American Journal of Psychiatry* was facilitated by the journal editor who also had financial ties to GSK. Thus, GSK was able to take an under-powered and non-informative trial with negative results and present it as a positive marketing vehicle for off-label promotion of paroxetine for bipolar depression. In addition to the commercial spin of paroxetine efficacy, important protocol-designated safety data were unreported that may have shown paroxetine to produce potentially harmful adverse events.

Introduction

Since the late 1990s, industry-sponsored ghostwriting for the medical literature has been exposed in a number of articles, plaintiffs' lawsuits and government investigations. Ghostwritten articles have been the main vehicle by which the pharmaceutical and medical device industries engage in the deliberate misrepresentation of efficacy and safety for product promotion and legal defense.[1,2]

In this regard, Sally K. Laden was one of the most prolific ghostwriters of the now-defunct medical communication company, Scientific Therapeutics Information (STI). A recently-posted document describes a marketing and publication plan by SmithKline Beecham, now GlaxoSmithKline (GSK) that commissioned STI to produce a manuscript for paroxetine treatment of panic disorder. In her correspondence with GSK managers, Laden wrote: "There are some data that no amount of spin will fix…."[3] On the other hand, substantial data appeared in numerous articles ghostwritten by Sally Laden such as the paroxetine 329 adolescent depression trial, the paroxetine 352 bipolar depression trial and a special issue of *Psychopharmacology Bulletin* "Advancing the Treatment of Mood and Anxiety Disorders: The First 10 Years' Experience with Paroxetine."[1,4-5] If Laden is acknowledged at all for having produced these manuscripts, she is mentioned only in the fine print, as in the case of study 329, for "editorial assistance."[6] These and other ghostwritten publications for GSK are listed among over 100 projects that GSK commissioned from STI in a document entitled, "Paxil-Funded Publications 1998 to Current," subpoenaed in 2008 in the case of *Burdick vs. GSK*. In addition to ghostwriting articles and letters to the editor, STI ghostwriters also prepared continuing medical education programs, conference posters, speaker training materials, and slide kits. For example, project #1305, Nemeroff study #352, describes the project as follows: "Write up of a clinical study. STI provided editorial assistance to primary author: writing/editing, coordinating materials needed for journal submission."[7] Another item on this document is project #1112, "Handbook: Psychopharmacology, […] developed for clinical practitioners; overall editor C. Nemeroff," which was

[i] Depression Research Unit, Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA (jamsterd@pennmedicine.upenn.edu)

[ii] Department of Philosophy, California State University, Northridge, CA, USA (leemon.mchenry@csun.edu)

Received: August 2019
Accepted for Publication: October 2019
Published Online: November 2019

published as a textbook entitled *Recognition and Treatment of Psychiatric Disorders: A Psychopharmacology Handbook for Primary Care* under the authorship of Charles B. Nemeroff and Alan F. Schatzberg. It is hotly disputed as to whether this textbook was ghostwritten by STI.[8]

"Editorial assistance" was the euphemistic term STI used to describe ghostwriting. If, however, it were merely a matter of editorial assistance in the technical write up of clinical trial results, there would be no issue with what has become common practice. The problem, as we will demonstrate below, is concealing corporate input that misrepresents the data for marketing purposes.

As a result of newly-publicized documents that provide significant evidence of the misreporting of clinical trial results, this case study is a follow-up to our previous article, "The paroxetine 352 bipolar trial: A study in medical ghostwriting," published in the *International Journal of Risk and Safety in Medicine*.[5] In that article we critically evaluated the paroxetine 352 bipolar trial in order to demonstrate how the published report conflated primary and secondary outcome analyses, turning negative results into positive results—with conclusions that could adversely affect patient health. In the current deconstruction article, we examined forty-two publicly-available documents produced by Scientific Therapeutics Information, Inc. from paroxetine litigation. Forty of the documents were produced as part of *In Re: Paxil, C.P. Ct. PA (On-Drug)* and concern the production of the report of paroxetine study 352. The remaining two documents were related to the publication plans for paroxetine and another clinical trial, paroxetine study 222 for panic disorder. The documents are posted on the Drug Industry Document Archive (DIDA) web site at the University of California San Francisco.

The Paroxetine 352 Study and its Findings

Study 352, entitled "Double-Blind, Placebo-Controlled Comparison of Imipramine and Paroxetine in the Treatment of Bipolar Depression," was published in the *American Journal of Psychiatry* in June 2001 under the byline of Nemeroff *et al*. (2001).[9] It was designed as an 18-site, 10-week, randomized, double-blind, placebo-controlled comparison of paroxetine versus imipramine in subjects with bipolar type I major depression unresponsive to lithium carbonate at therapeutic plasma lithium levels. It had a projected study duration of 2 years. Its protocol-designated objective was to compare the efficacy and safety of paroxetine and imipramine to placebo in the treatment of bipolar depression in subjects stabilized on lithium therapy.

The primary efficacy measures were the change from baseline Hamilton Rating Scale for Depression (HRSD) total score, and the change from baseline in the Clinical Global Impression Severity (CGI/S) score for paroxetine versus placebo and for imipramine versus placebo.[9-11] The comparison of primary interest was paroxetine versus placebo irrespective of baseline lithium level stratification. Protocol-stipulated secondary outcomes included the proportion of subjects with a final HRSD score ≤ 7 or a final CGI/S score ≤ 2 as well as the proportion of subjects with adverse events, premature treatment discontinuation and manic or hypomanic symptoms as measured by the DSM-III-R Mania/Hypomania Assessment and the Young Mania Rating Scale (YMRS).[12] Analyses were to be performed on the entire subject population and on subjects who experienced a manic or hypomanic episode (versus those who did not). The YMRS measure was to be used to assess severity of manic and hypomanic symptoms across treatment groups, and the relationship between change in YMRS scores and HRSD scores was to be examined.

The study population consisted of outpatient subjects ≥ 18 years old with a history of ≥ one prior manic or depressive episode within the preceding 5 years. The original protocol called for a sample size of 62 subjects per treatment condition (i.e., totaling 186 subjects). However, due to poor subject enrollment, only 117 subjects were recruited into the study, resulting in a final sample size distribution of paroxetine (n=35), imipramine (n=39), and placebo (n=43). As a result, the paroxetine 352 study was a non-informative trial with insufficient statistical power to show anything other than inconclusive results. There was no evidence of paroxetine or imipramine efficacy versus placebo in bipolar major depression, and the presentation of safety data were insufficient to draw any clinically meaningful conclusions. Despite the protocol stipulations, Nemeroff *et al*. failed to mention that the YMRS rating was employed as an outcome measure in the 352 study and most other manic and hypomanic safety ratings obtained with the YMRS measure were omitted from the published article. As a result, the published article inaccurately reported that there was no evidence of any paroxetine-induced manic and/or hypomanic symptoms in bipolar major depressive disorder while the suppression of safety data from the YMRS outcome measure hid the presence of possible paroxetine-induced harm.

Complaint of Research Misconduct in the 352 Study

In July 2011, a Complaint of Scientific Misconduct was filed with the Office of Research Integrity (ORI) of the Department of Health and Human Services against Dwight L. Evans, Laszlo Gyulai, Charles B. Nemeroff, Gary S. Sachs, Charles L. Bowden *et al*.[13] As part of the formal adjudication process of the ORI misconduct case, the University of Pennsylvania organized a formal ORI Inquiry Committee comprised of three professors from the School of Medicine to investigate the allegations of misconduct.

According to the complaint, the paroxetine 352 article was ghostwritten by employees of GSK and STI. It was published by the *American Journal of Psychiatry* in June 2001 under the author byline of Nemeroff *et al*. without acknowledging the role of GSK or the STI ghostwriters. The complaint alleged: [14]

(1) "that Dr. Evans and Dr. Gyulai allowed their names to be appended to a manuscript drafted by a medical communications company and, thereby, Dr. Evans and Dr. Gyulai were not legitimate authors of the manuscript" (p.3);

(2) "that the manuscript was 'ghostwritten' by STI and that the authors of the published manuscript failed to appropriately acknowledge STI's contribution" (p.3);

(3) "the preliminary drafts of the study 352 article were conceptualized and drafted by STI and not by any of the named authors" (p.5);

(4) "Dr. Evans' contribution to the preparation of the manuscript was limited to his commenting on, and approving, STI and GSK ghostwritten drafts of a manuscript on which he was designated as second author, and of which he had no direct knowledge of the accuracy of the data analyses, data interpretation (i.e., the inclusion or exclusion of particular data analyses related to safety and efficacy), or the accuracy, or the information that was written in the manuscript (by the ghostwriters)" (p.5);

(5) "Dr. Evans has been engaged in lending his name to ghostwritten articles with the same 'ghosts' (i. e., Sally Laden) at STI beginning in 1997 to at least 2003" (p.6);

(6) "[Dr. Amsterdam] was intentionally left off from the review of the data and the drafting of the manuscript because the study sponsor, GSK, and the other 'authors' knew Dr. Amsterdam's professional ethics would not allow him to lend his name to a ghostwritten work, and most importantly, his morals would not allow the alteration and manipulation of data and would not allow the other 'authors' to turn a failed study into an undisclosed promotional marketing manuscript for the sponsor" (p.7);

(7) "despite STI's significant role in the preparation and drafting of the manuscript, the final published article makes *no mention* of STI's role in the article and does not mention that three of its authors, Ivan P. Gergel, M.D., M.B.A., Rosemary Oakes, M.S. and Cornelius Pitts, RPh. are GSK employees." (p. 8);

(8) "Dr. Nemeroff and Dr. Evans had very little, if any, direct input into the daily conduct of the 352 study, and certainly not enough to warrant being listed as the first and second authors on a manuscript published in one of the world's leading medical journals. Rather, their positions as authors on the manuscript were solely determined by GSK for the purpose of appending the names of 'key opinion leaders' to the manuscript for marketing and commercial promotion of paroxetine" (p.10-11).

In response to these allegations of ghostwriting and plagiarism, the University of Pennsylvania and the academic authors insisted that there was no involvement with any ghostwriters and that the allegations of research misconduct were unfounded.[14] The academic authors insisted that they had personally participated in the drafting of the first 352 manuscript. For example, an email press release by the University's spokesperson, Susan Phillips, indicated that an inquiry into the misconduct allegations "clearly concluded that this was not a case of ghostwriting or plagiarism."[14] Likewise, in a public press statement, Dr. Evans (who was second author on the published article and Chair of the Department of Psychiatry at Penn) wrote: "After a thorough review, the inquiry concluded that each and every allegation lacked substance and credibility."[14] In a similar fashion, Dr. Nemeroff (who was first author on the published article and Chair of Psychiatry at the University of Miami) reportedly told *Nature Magazine* that, while he was aware of STI's involvement in the preparation of the 352 manuscript, "All Sally Laden did was help collate all the different authors' comments and help with references. We wrote the paper."[14] In contrast to Nemeroff's assertion, Dr. Gary Sachs (fourth author on the manuscript and Professor of Psychiatry at Harvard) told the *Boston Globe* that he was "perplexed" by the allegations of ghostwriting and wrote in an email to the reporter: "When the data became available, I went to Philadelphia to help Dr. Gyulai [third author on the published article and Associate Professor of Psychiatry at Penn] draft the manuscript. We started with a blank page."[14] Sachs also told *Science Insider* that he did not know that STI ghostwriters were involved with the manuscript preparation. Finally, Dr. Charles Bowden (fifth author on the published manuscript and Professor of Psychiatry at the University of Texas Health Science Center) stated: "I never had any sense that the manuscript was 'ghostwritten.'"[14]

Deconstruction of the Ghostwritten 352 Manuscript

On March 13, 1997, the first draft of the 352 manuscript, prepared by Grace Johnson and Sally Laden of STI, was sent to Dr. Muriel Young, Medical Director at GSK. The cover page of the manuscript contained no author names.[15] By early April 1997, GSK senior managers leading the project appeared to have made certain alterations to the content of the first manuscript draft and these revisions were now incorporated into the second manuscript draft. An email from Sally Laden to Dr. Young at GSK dated April 4, 1997, stated:[16]

> We are pleased to enclosed Draft II of the manuscript "A Double-Blind, Placebo-Controlled Comparison of Imipramine and Paroxetine in the Treatment of Bipolar Depression." The manuscript has been modified based on your comments and those of Ivan Gergel, Cornelius Pitts, and Rosemary Oakes. Please note that some of Dr. Gergel's comments in the results section have been addressed in the discussion. We have also included one set of references as per Dr. Gergel's request.
>
> Dr. Gergel has confirmed that the *American Journal of Psychiatry* is the target journal for publication. The manuscript will be styled according to the journal specifications for the submission draft.
>
> We will contact the named authors (e.g., Laszlo Gyulai, Gary Sachs) once we receive your approval. At that time, Draft II of the manuscript will be sent immediately to Drs. Gyulai and Sachs for their review.
>
> We look forward to receiving your comments on Draft II of the manuscript by April 25, 1997. We will incorporate your comments and those of the named authors and submit Draft III for your review on May 2, 1997.

This document indicates that the second manuscript draft still had no author names listed on its cover page and that the manuscript contained only GSK-directed revisions to draft #1. The Laden email was only sent to GSK managers for their review. This document also indicates that Drs. Gyulai (of Penn) and Sachs (of Harvard) were likely designated by GSK as the original lead authors on the 352 manuscript. Thus, the first two manuscript drafts were entirely ghostwritten under the direction of GSK, without any contribution from named academic authors. Finally, the document suggests that none of the academic authors ever saw manuscript draft #1 or #2, until draft #2 revisions were completed and verified by GSK.

In a subsequent May 1997 correspondence, Cornelius Pitts and Ivan Gergel of GSK provided to the STI ghostwriters their hand-written revisions to manuscript draft #2. This revised manuscript draft indicates that the presentation of the *post hoc* "finding" of positive paroxetine efficacy in subjects with low baseline lithium levels was solidified into the manuscript text as if it were the primary finding of the study. In fact, this was the *only* "positive" statistical analysis for paroxetine in the entire study and, according to the study protocol, was unnecessary and should not have been performed. This revised draft now contained the omission of any weekly mania rating analyses or mania and hypomania symptom ratings (which were specifically stipulated in the study protocol). Furthermore, without any supporting evidence, this revised draft now indicated an absence of paroxetine-induced sexual side effects and an absence of paroxetine-induced mania. The second draft solidified the total sample size of the study at 117 subjects as if this were the protocol stipulated sample size. In this regard, all information on sample size estimates was removed from this draft, and the final power estimate provided in this draft was finessed to comport with the truncated study enrollment of 117 subjects (rather than the projected 186 subjects). Finally, GSK directed the ghostwriters to revise the discussion section of the draft to lend a favorable commercial spin to paroxetine by having them indicate that paroxetine is as efficacious as imipramine when, in fact, neither of the medications was superior to placebo in any of the outcome analyses.[17] Thus, for example, while draft #2 reads: "Paroxetine and imipramine are comparable in efficacy for the treatment of bipolar depression in patients maintained on low-lithium levels," the GSK hand-written revision reads: "Paroxetine and imipramine are comparable in efficacy for the treatment of bipolar depression showing statistically significant superiority to placebo in patients maintained on low-lithium levels."[17] STI-edited iterations of these GSK-directed revisions were all eventually incorporated into the published article (p.910).[9]

Honorary Authorship of Academic Key Opinion Leaders

A May 19, 1997 fax from Dr. Young at GSK to Grace Johnson at STI provides a revealing picture of how the ghostwriters appended the names of the GSK-designated "authors" to the developing 352 manuscript. More than 2 months after the first draft of the manuscript was produced for GSK, Dr. Young directed the STI ghostwriter to add the following author to draft #3:[18]

> I would like you to add the following two names as authors on the paper:
>
> Charles Nemeroff, M.D.
> Emory Clinic
> 1365 Clifton Road, N.E.
> Room 5312, Psychiatry Dept.
> Atlanta, GA 30322
>
> Dwight Evans, M.D.
> University of Florida Health Science Center
> Shands Hospital 1600 S.W. Archer Road 11th Floor
> Gainsville, EL 32610-0486

Dr. Young then goes on to direct the ghostwriter how to arrange the academic authors' names in the byline of draft #3 and which GSK employees to add as authors. She wrote:[18]

> The lead author should be Lazslo Gyulai, then Charles Nemeroff, Gary Sachs & Dwight Evans, in that order. The other authors are Muriel L. Young, M.D., Ivan P. Gergel, Cornelius Pitts, & William Bushnell. Please send a copy of the latest draft of the manuscript to Drs. Nemeroff and Evans.

Thus, this document shows that GSK directed the ghostwriters to add Nemeroff and Evans as authors to the manuscript after two drafts had already been produced and revised by GSK. This GSK-designated order of author names would eventually change over time from draft #3 to the final manuscript draft. For example, it is noteworthy that, at this stage in the manuscript development, there is no mention of Dr. Charles Boden as an "author" (although his name will appear on later manuscript drafts and on the published article). Moreover, with the exception of Drs. Gyulai and Sachs (who were initially designated by GSK as lead "authors" after draft #2 was completed), the remaining academic "authors" only had their names appended to the author byline after draft #3 was completed. To this end, on May 20, 1997, Grace Johnson of STI confirmed to Dr. Young that letters of invitation had indeed been sent to Drs. Nemeroff and Evans, inviting them to be listed as "authors" on the 352 manuscript. For example, Grace Johnson wrote:[19]

> Dear Dr. Evans:
>
> By way of introduction, my name is Grace Johnson, and I am an editor at Scientific Therapeutics Information, Inc (STI). STI is working with Dr. Muriel Young at SmithKline Beecham (SB) Pharmaceuticals to develop a manuscript "A Double-Blind, Placebo-Controlled Comparison of Imipramine and Paroxetine in the Treatment of Bipolar Depression."
>
> We are pleased to invite you to participate as an author on this article. Enclosed please find Draft II of the manuscript for your review. You may mark your comments directly on the manuscript. Please return your comments on Draft II of the manuscript by June 12, 1997. We will incorporate your comments and those of the other reviewers and submit Draft III for your review. Following the approval of Draft III, the manuscript will be styled for submission to the *American Journal of Psychiatry*.
>
> Please do not hesitate to contact us with any questions or comments during your review. We look forward to working with you on this project and to submitting the manuscript for publication.

Thus, despite their later, public assertions that the 352 article was not ghostwritten, it is clear to us that Dr. Evans and Dr. Nemeroff were definitely aware of the existence of STI's involvement in the production of the manuscript. It is also clear to us that their names were being appended as "honorary" or "guest" authors on the third draft of the developing manuscript to which they had made no author contributions.

A subsequent, internal June 6, 1997 STI memorandum from Grace Johnson to her STI associates summarized the current status of the 352 manuscript project:[20]

> D[raft] 2 was sent to SB and Drs. Gyulai and Sachs on April 4, 1997. SB added to more authors (Dr. Nemeroff and Evans) on May 20, 1997 [*sic*]. To date, we have received comments from Dr, Young, Neil Pitts, and William Bushnell. Dr. Gyulai requested SB to complete additional statistical analysis on the data. So, D[raft] 2 of the manuscript is currently on hold until the analysis is complete. The proposal states four drafts. However, with so many reviewers and additional statistical analysis on D[raft] 2, another draft might be needed. (This may depend on if the statistical analysis changes the results section of the study.)

At this point, there were a total of 4 GSK-designated academic authors, i.e., Drs. Gyulai, Sachs, Nemeroff, and Evans, and 4 GSK employees. Dr. Bowden's name had not yet been added to the author byline. It was also noteworthy that Dr. Gyulai's request for additional *post hoc* statistical analyses by GSK would result in greater production costs to the sponsor (see below). Finally, this internal memorandum noted that GSK had authorized STI to develop 25 PowerPoint slides describing the results of the 352 study (presumably for marketing and educational purposes). This project was to be completed in May 1997, ahead of the publication of the 352 article. By December of 1997, it appeared that the lead author, Dr. Gyulai, had fallen into disfavor with GSK for his requisition of additional statistical analyses, which led to a delay in the completion of the 352 project.

In a December 15, 1997 letter from Sally Laden to Dr. Nemeroff, Laden indicated that Charles Bowden had now been designated by GSK as a 5th academic author and that Dr. Gergel of GSK had advised STI that he, Dr. Nemeroff, would heretofore be designated by GSK as lead author on the 352 project. Laden wrote:[21]

> Ivan Gergel asked me to send you the enclosed author comments on the Paxil bipolar disorder study report. STI progressed to the Draft II stage at which point Dr Gyulai has held onto the paper without response to us.
>
> Enclosed are comments from yourself, Dwight Evans, and Charles Bowden.
>
> Ivan mentioned that you will be leading development of this manuscript in the future, but that arrangements with Dr Gyulai have not yet been finalized. Once I learn more from Ivan, I will be in touch with you to discuss our next steps.
>
> Thanks for understanding this process. Please don't hesitate to call me if you have questions or require additional information.

Four months after Laden informed Nemeroff that he would assume the lead on the 352 project, Laden sent an April 3, 1998 proposal to Dr. Gergel at GSK to formalize the designation of Nemeroff to restart the stalled 352 project. The proposal contained a detailed explanation for an additional $10,000 cost of completing the 352 article. It further detailed the cost for every step of the ghostwriting,

review and submission of the manuscript. The proposal stated:[22]

> STI will develop up to four drafts: Draft I is the initial draft that we will receive from Dr Nemeroff and which we will edit and follow his direction. This will be reviewed by the sponsor and Dr Nemeroff, Comments on Draft I will be incorporated into Draft II, which will be sent to the same reviewers for review and critique. Draft III is the pre-journal submission draft and will be sent to the author and sponsor for final review and approval prior to developing the journal submission package. Draft IV is the journal submission draft prepared for the journal, for which art work will be professionally drawn and the manuscript styled according to the selected journal.

It should be noted also that page 7 of this April 3rd proposal indicates in the "Time and Events" section that SB, i.e., GSK, will receive draft 1 from Nemeroff one month after STI receives the manuscript. This statement, however, contradicts earlier STI documents (cited above). In this regard, draft #1 had already been produced by STI on March 13, 1997.[16,17] Thus, irrespective of whether Laden's April 3, 1998 designation of "draft #1" is the same as the original draft #3, a comparison of the two STI manuscripts makes it immediately apparent that both drafts contain the same information that is present in the original STI draft #1 from March 13, 1997. From our reading of the documents, the most plausible interpretation is that the April 3, 1998 STI contract from Laden to Gergel did not begin afresh with a new 352 draft #1.

Role of the Journal and the Peer-Review Process

On February 24, 1999, Sally Laden (under Dr. Nemeroff's name) began the submission process of the 352 manuscript to the *American Journal of Psychiatry*. Laden wrote to each GSK-designated author to obtain their signed copyright release to the journal. The "Manuscript Submission Approval and Copyright Transfer" reads:[23]

> I have been sufficiently involved in this work to take public responsibility for its validity and final presentation as an original publication. I can and will provide documentation of my work upon reasonable request and I have fulfilled the obligations for full disclosure and authorship as described by the *American Journal of Psychiatry*.

In a subsequent communication from Laden to Nemeroff dated March 5, 1999, Laden informed Nemeroff that: [24]

> After I sent what I thought was the final version (SB had given it their blessing) to all authors for their sign-off, Dr Gyulai pointed out that the ad-hoc analysis of high vs low HAMD patients excluded those patients with thyroid changes. SB quickly re-ran the data and found that the significant differences between imipramine and placebo and the trends for paroxetine were lost. Thus, we deleted Figure 2 and deleted one sentence from the text. Because this wasn't the main point of the paper, I'm hoping that this change is OK. I'm assuming that it is and am sending the submission package to you today.

Thus, it appears that Laden unilaterally revised the manuscript (i.e., without any consultation from the academic or GSK authors) to avoid reporting the newly-analyzed, *post hoc* "negative" finding. Thus, if Laden had not made this unilateral editorial decision, there may have been no "positive" analyses of any sort to report on paroxetine.

An email thread between Nemeroff and Laden dated July 13, 1999 that was forwarded to Cornelius Pitts at GSK provided information on the peer-review status of the 352 manuscript by the *American Journal of Psychiatry*. Laden wrote: [25]

> I'm reviewing my list of projects with SB and have a few questions. 1. Have you heard from Am J Psychiatry about the bipolar paper?...

Nemeroff responds:[25]

> I am in Zurich as I respond at midnight visiting with Roche. The answers are: 1. Jack Gorman [Deputy Editor of the American Journal of Psychiatry] told me that the Bipolar paper has come back from review and will be accepted after revisions. I haven't received the reviews yet but you will be the first to know, I promise….We love working with you.

A subsequent September 13, 1999 email from Laden to Nemeroff (copied to Pitts at GSK) provided additional information on the status of the peer-review process. Apparently, the peer-reviewed manuscript was returned from the *American Journal of Psychiatry* and was back on the desk of Sally Laden at STI. Laden informed Nemeroff that extensive revisions were requested by the reviewers and that many of the comments involved statistical issues that she could not address, as she (i.e., Laden) revised the peer-reviewed manuscript. As a result, Laden sent the partly-revised manuscript to Pitts at GSK to address the requested statistical revisions with the GSK statistician. Laden wrote to Nemeroff: [26]

> Last week, I finished with my first cut at the Am J Psych reviewers' extensive comments to this paper. Many of their comments were statistical in nature, which only the SB statistician can answer. Neil Pitts has it now and I'll stay in touch with him about the progress of their review.

A subsequent email exchange, between Laden and Pitts, dated March 15, 2000 provides more insight into the STI/GSK revisions that were made to the 352 manuscript. The revisions suggested by Pitts to Laden were memorialized in an email "note-to-file" made by Sally Laden to herself: [27]

> Note to self:
>
> make these changes to the bipolar manuscript

> 1. emphasize that only 5 of 8 patients were in the efficacy population and this didn't come through in the manuscript revision.
>
> 2. tone it down and the efficacy of those 5 patients who had efficacy change scores was similar to the overall study.

In a subsequent April 5, 2000 email, Laden provides Nemeroff with a further update on the revisions of the 352 manuscript: [28]

> The bipolar manuscript has been revised [by GSK] to include the statistical reanalysis, and Rocco Zanninelli and Neil Pitts have signed off on it. It is being fedexed to you tomorrow and you will receive it on Wednesday. If you approve the manuscript, I will prepare a brief response to the journal so that you can resubmit it.... Again, upon your approval to proceed, I'll have a submission package prepared for you to submit the manuscripts to the journal.

This email from Laden to Nemeroff, together with the September 13th email exchange above, indicate to us that Nemeroff and the other academic authors had little or no scientific input into the original or revised 352 data analyses and that virtually all of the revisions were made by Sally Laden and GSK employees i.e., Drs. Zaninelli, Pitts, Oakes, and Gergel. Sally Laden, and not Nemeroff, wrote the revision cover letter to the editor of the *American Journal of Psychiatry*) (i.e., Jack Gorman). In a somewhat curious addendum to this email exchange, Nemeroff wrote to Laden: [28]

> Ivan Gergel, MD who was with SKB and was a coauthor [on manuscript draft #3] must be reinserted as a coauthor. I will not compromise on this as he was involved in the early drafts and this is simply an ethical issue. Dwight and other authors agree about this. We must take the high road here.

In an email dated May 11, 2000, Laden wrote to Dr. Evans and the other academic authors indicating that she needed their approval of the revised 352 manuscript that was prepared for Nemeroff. Laden advised the authors of the revision process:[29]

> Significant re-analysis has been accomplished and the revisions have been made and approved by Dr Nemeroff, Rocco Zaninelli, MD (of SB), and Neil Pitts, RPh (of SB). Dr Nemeroff will be submitting the revised manuscript to the journal early next week. A copy of the revised manuscript (note that new references are not yet renumbered) and Dr Nemeroff s point-by-point rebuttal to the journal are attached in WORD files below.

Evans responded to Laden:[29]

> Dear Sally,
>
> Masterful! No additional changes recommended.
>
> Dwight

An email dated June 12, 2000, indicated that further difficulties had arisen in publishing the 352 manuscript. It appears that the *American Journal of Psychiatry* had sent the revised manuscript to an additional, independent statistical reviewer who rejected the revised manuscript and instead recommended additional analyses, revisions and the elimination of all commercial bias. Although the official letter from Gorman to Nemeroff, printed on *American Journal of Psychiatry* letterhead, was not sent to Nemeroff until June 22, 2000, concerns about the manuscript were already being expressed by the STI ghostwriter and GSK. In this regard, Laden wrote to Nemeroff on June 12th: [30]

> Thanks for copying me on the email from Jack [Gorman] about Am J Psychiatry's review. I've passed this along to SB and await their reaction. What are your feelings about this?

Nemeroff responded to Laden:[30]

> Jack is clearly willing to fight for the paper. The statistician's point is well taken, namely the secondary hypothesis testing. I think we should see if the SB statisticians are able to respond adequately to the criticisms and moreover if we can tone down the perceived commercial bias so Jack can pull it over the line. If not, we can publish it relatively quickly in Depression and Anxiety. I can use the AJP reviews from the first review.

Laden replied to Nemeroff that she will pass the information about the need for additional revisions and the assistance from Jack Gorman to the GSK managers. Gorman had significant financial ties to GSK as a recognized key opinion leader, member of GSK's psychiatry advisory board, speakers' bureau and frequent spokesperson for paroxetine (Paxil) on television programs and advertisements.[31]

Despite Nemeroff's reassurance to Laden, it appears that GSK was, nonetheless, increasingly frustrated with the additional time, cost, and uncertainty of getting the 352 manuscript published in the *American Journal of Psychiatry*. To wit, an internal email among STI employees, dated June 13, 2000, revealed GSK's dissatisfaction with STI's ghost management of the 352 project. Marion Philips and Sally Laden describe to STI President (John Romankiewicz) the extent of GSK's dissatisfaction:[32]

> Marion - this project (the bipolar paroxetine manuscript) has come into trouble with the client [i.e., GSK]. It has been a protracted process - we are currently at Draft XI, if my calculations are correct. They are disgusted with the author (Charlie Nemeroff) and with the fact that they have to pay an agency [i.e., STI] extra to get it done. Remember, this is the client (Rocco Zaninelli [Medical Director at GSK] who refuses to work with any outside medical writers, including STI. A brief history is this:
>
> 1. Submitted to Am J Psychiatry,

2. Journal sent back significant revisions, but no promise of acceptance

3. Many drafts later, the revised version (draft X or XI) was resubmitted

4. Yesterday, journal responded that they might consider it, but significant revision was still needed.

5. Charlie Nemeroff feels that with more revision we might get it in the Am-J Psychiatry, but that we could probably get it published in his journal (Depression & Anxiety or the British J Psychiatry)

I advised client of this immediately. They responded by saying that they don't care where it is published, they just don't want to be charged extra. They want to turn this [i.e., the production of the 352 manuscript] over to Charlie so that he can finish it up.

Meanwhile, behind the scenes, a private email from Gorman to Nemeroff, dated June 12, 2000, detailed how Gorman could facilitate the revision and publication of the rejected 352 manuscript. Gorman indicates that he was not quick enough in forestalling the statistical reviewer from rejecting the 352 manuscript; but would nonetheless get it published, with GSK's direct assistance. Gorman wrote:[33]

I got the bipolar depression paper back and wanted to give you a heads up. The revision went automatically to the statistical reviewer before coming to me, which is the usual custom on resubmitted manuscripts. The statistical reviewer adamantly recommends rejection. I read it again carefully and I still think the paper is important as one of the only if not the only one of its kind. However, journals are now under fierce scrutiny about drug company sponsored studies, thanks in large part to Marcia Angel's recent stuff in the NEJM.

The paper still has a very biased tinge to it. Look, the fact is that the study was beautifully designed and the results are important. But in the high lithium level group the two active drugs did no better than placebo and it is a legitimate conclusion that raising the lithium level alone may be sufficient to treat bipolar depression. I think you would want to make that point much more strongly.

Second, using the main analyses, the ITT for all groups, there is no statistically significant drug versus placebo effect. The journal will never allow 'numerical superiority' as a standard. So the second conclusion is that for the study as a whole, drugs didn't do better than placebo.

In a secondary analysis you can show that if you don't want to raise lithium level, there is a completer finding that imipramine and paroxetine are better than placebo, but it must be labelled as a very post hoc observation.

Finally, we all know that paroxetine is better tolerated than imipramine but I don't see it in this data set. None of the comparisons for adverse side effects or drop outs is statistically significant, [and] the claim that paroxetine is better on that basis than imipramine is not warranted.

You'll get all this back in the form of an invitation to revise again. Even if I bend, no way NCA [Nancy C. Andreason, Chief Editor] would allow it in its present state so it is moot for me to go ahead. I think you might want to go over it again with the SKB people and let them know that the study just doesn't support much in the way of a claim that paroxetine worked particularly well here compared to other interventions.

If I've missed something, of course let me know. I want to be helpful and this was a lot of work and a great study.

In follow up to this June 12th, back channel email from Gorman to Nemeroff, Laden sent a fax correspondence on July 5, 2000 to Drs. Raj Kumar and Rocco Zaninelli of GSK to reassure them that the 352 manuscript would most probably be accepted for publication by the journal editor, if a few minor cosmetic revisions were made to the manuscript. However, Laden also recognized that GSK may not want to expend additional time and money on publishing the 352 manuscript in the *American Journal of Psychiatry*; in which case it will definitely be accepted in Nemeroff's *Depression & Anxiety* journal.[34]

It appears that Nemeroff was probably successful in gaining Dr. Kumar's approval to have GSK continue working with STI and Gorman. As a result, Dr. Kumar agreed to allow STI to move forward with another revision. This resulted in a July 26, 2000 letter in which Laden writes to Kumar:[35]

At the request of Charlie Nemeroff, I am enclosing the most recent correspondence from *Am J Psychiatry* related to this manuscript. Charlie tells me that you are willing to continue revising this manuscript and submit it a third time to the journal. According to Charlie, this will be accomplished as follows:

1. Rosemary Oakes will address the statistical reviewer's comments and make changes to the manuscript.

2. When done, she will send these changes to STI to be incorporated into the manuscript.

3. STI will send the revised manuscript to Charlie, who will revamp it and remove the 'marketing' tone to which the journal objected.

4. Charlie will either submit the manuscript directly or return it to STI to be restyled for resubmission to the journal.

Enclosed please find a copy of the manuscript plus the journal's reviews and correspondence.

I trust that this is an accurate summary of the agreement made to finish this project. If not, kindly contact me directly.

A subsequent October 31, 2000 email from Laden to Nemeroff details the revisions that were made by Laden and GSK to the 352 manuscript. Laden indicates that the revised

manuscript fulfills all of Gorman's June 22, 2000 officially-recommended revisions. Laden wrote to Nemeroff:[36]

> The changes consist of the addition of more statistical detail in Table 2 and greatly toned down [the commercial] text in the Abstract, Results, and Discussion about the findings.

Laden notes that:[36]

> … this version is sufficiently noncommercial and fulfills the directives outlined by the journal.

Finally, Laden indicates to Nemeroff that if he approves the Gorman and GSK-recommended revisions, she will prepare a submission packet for Nemeroff containing:[36]

> a detailed response to the reviewers, 4 copies of the manuscript and cover letter, a new glossy for the figure, and all files on a disk.

On November 11, 2000, Laden ghostwrote a cover letter (in Nemeroff's name) to Jack Gorman at the *American Journal of Psychiatry*, detailing the second revision to the 352 manuscript.[37] This letter was then placed on Emory letterhead for submission to the journal via Nemeroff's office staff. The cover letter states that the report of the 352 trial:[37]

> represents the largest placebo-controlled trial of bipolar depression ever conducted [*sic*], and contains important and novel information to the field.

It is noteworthy, however, that there is no mention of the study being prematurely terminated for insufficient subject enrollment or that it had insufficient power to test any of the primary or secondary study aims. It also states that the study outcomes are unique, when, in fact, the GSK authors, journal reviewers and journal editor noted that the study was a negative trial.[24,33]

While the peer-review and revision process of the 352 manuscript finally resulted in its acceptance for publication by the *American Journal of Psychiatry* on January 6, 2001, we would note that this process (such as it was) failed to adhere to the International Committee of Medical Journal Editors (ICMJE) policy regarding authorship, according to which both honorary authorship and ghost authorship are considered forms of misconduct. The *American Journal of Psychiatry* policy in place at the time of the submission of the Nemeroff *et al*. manuscript stated:[38]

> All persons designated as authors should qualify for authorship. Each author should have participated sufficiently in the work to take public responsibility for the content.
>
> Authorship credit should be based on substantial contributions to: 1) conception and design or analysis and interpretation of data, and 2) drafting the article or revising it critically for important intellectual content, and on 3) final approval of the version to be published. Conditions 1, 2, and 3 must all be met. Participation solely in the acquisition of funding or the collection of data does not justify authorship. General supervision of the research group is also not sufficient. Any part of an article critical to its main conclusions must be the responsibility of at least one author.
>
> Only those with key responsibility for the material in the article should be listed as authors; others contributing to the work should be recognized in an Acknowledgement. Editors will require authors to justify the assignment of authorship.

On September 22, 2016, the current authors wrote to the *American Journal of Psychiatry Chief* editor, Robert Freedman, to request retraction of the Nemeroff *et al*. article on the basis of the violation of journal policy and the misrepresentation of study 352 results. Dr. Freedman did not respond, and thus the scientific record remains uncorrected. (The letter from the current authors is published as a supplemental file to this article).

Discussion

The paroxetine 352 study was a non-informative trial with insufficient statistical power to show anything but inconclusive results. There was no evidence of any paroxetine (or imipramine) efficacy relative to placebo in bipolar major depression and the suppression of safety data hid the presence of paroxetine-induced harm.

The misrepresentation of the 352 study as a published report was facilitated by STI ghostwriters working in conjunction with GSK employees and the editor of the *American Journal of Psychiatry*. Given the documentary evidence presented above, and despite their uniform denials to the Penn ORI Inquiry Committee, it is difficult to see how the academic authors could not have been aware that a medical writing firm was involved in the manuscript production. Moreover, most of the academic authors apparently misled the *American Journal of Psychiatry* by appending their signature to the "Manuscript Submission Approval and Copyright Transfer" that claimed they were sufficiently involved in the work to take public responsibility for its validity and final presentation. These academic "authors" failed to adhere to the journal's authorship criteria and they violated the policy by failing to disclose the role of STI and GSK in the development of multiple manuscript drafts.

The only mention in the published Nemeroff *et al.* article of drug company support is the rather misleading statement, "Supported by NIMH grant MH-51761 and a grant from GlaxoSmith-Kline," suggesting that the primary financial support for the 352 study was provided by a NIMH supported grant with secondary support from GSK.[9] This, of course, was not the case.

From our deconstruction of the documentary evidence, it appears that none of the academic authors had access to, or reviewed, the actual patient-level data or the repeated GSK-produced data analyses and it is certainly false that they had complete freedom to direct its analysis and its reporting without influence from the sponsor. We would also note that the ethical protestation by Nemeroff, Evans and the other academic authors regarding the guest authorship of Ivan Gergel [after his departure from GSK to Forest Laboratories, Inc.] is ironic given the fact that none of the academic "authors" had any involvement in the 352 study and that they were all designated by GSK for inclusion in the author byline. It is also obvious to us that the journal editor, Jack Gorman, took unethical and unprofessional editorial license with the journal peer-review process in order to facilitate publication of the 352 manuscript. Gorman also failed to have the 352 authors justify the assignment of authorship. In our view, the authorship policy of the *American Journal of Psychiatry* is largely window dressing common for journals of the sort that publish industry-sponsored trials.

The ORI Inquiry Committee of the University of Pennsylvania ignored the allegations of research misconduct in the performance of the 352 study and, instead, focused on the ghostwriting aspect of the published article. In this regard, Penn failed to examine the available evidence in the case and ignored the most egregious evidence of misconduct, i.e., the newly-posted STI documents. This omission allowed the University to avoid public embarrassment and to clear their professors of any wrongdoing. Thus, instead of punishing their professors for academic misconduct, the University whitewashed the presence of any misconduct within its psychiatry faculty. Penn's conclusion of no wrongdoing by its professors provoked such a public outcry of indignation that the issue of ghostwriting at the University of Pennsylvania reached the desk of the President of the United States.[39]

Conclusion

Because ghostwriting is designed to evade detection and is only revealed as a result of litigation or government inquiries, it is therefore imperative to document the cases in which ghostwriting has facilitated misrepresentation of clinical trial results. The integrity of science depends on the trust placed in individual clinicians and researchers and in the peer-review system which is the foundation of a reliable body of knowledge. When academic researchers allow their names to appear on ghostwritten articles, they betray this basic ethical responsibility and are guilty of academic dishonesty. Medical journal editors are entrusted with significant power as gatekeepers of the scientific record. They also bear the responsibility to ensure that the journals are not publishing manuscripts guilty of fraud, fabrication and plagiarism. Ghostwriting is a serious problem because it is a dishonest attribution of the origin of the manuscript, it disguises marketing and public relations objectives of for-profit companies as science, conceals conflicts of interest of named "authors" on manuscripts, misrepresents the results of scientific testing, and, most importantly, has contributed to fatal consequences in cases in which the safety of drugs are misreported.

Limitations

The authors warrant that findings have been reported fairly and non-selectively. It is, however, always possible that there are gaps in the record due to inadequate response to discovery in litigation. Since the complete deconstruction of each STI document could not be exhaustively undertaken in the course of this review, we recommend that readers examine each STI document that is posted on the Drug Industry Document Archive (DIDA) web-site at the University of California San Francisco (*see* https://www.industrydocuments.ucsf.edu/drug/collections/paxil-litigation-documents/).

Acknowledgments and Disclosure:

This research was not supported by any federal, corporate, or private funding agency or grant. The authors thank Ronald Goldman, Esq. and Michael Baum, Esq., for legal review of the manuscript, the Drug Industry Document Archives (DIDA) for posting the paroxetine 352 study documents on their website, two anonymous reviewers from the journal and Tess Bird for helpful suggestions to the final draft. The views expressed herein are those of the authors alone and not those of any other person, firm or entity. Dr. Amsterdam received legal support from Baum, Hedlund, Aristei & Goldman of Los Angeles, California for his Complaint of Scientific Misconduct against Dwight L. Evans, Laszlo Gyulai, Charles Nemeroff, Gary S. Sachs and Charles L. Bowden provided to the United States Department of Health and Human Services Office of Research Integrity in July 2011. Dr. McHenry has been a research consultant for Baum, Hedlund, Aristei & Goldman since 2003.

References:

1. McHenry L. Of sophists and spin-doctors: Industry-sponsored ghostwriting and the crisis of academic medicine. *Mens Sana Monographs*.2010;8: 129-145

2. Sismondo S. Ghost management: How much of the medical literature is shaped behind the scenes by the pharmaceutical industry? *PLoS Medicine*. 2007;4(9): e286.

3. STI. Email from Sally Laden to Daniel Burnham. 2000 Dec 14. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=snxl0228

4. McHenry L, Jureidini J. Industry-sponsored ghostwriting in clinical trial reporting: A case study. *Accountability in Research*. 2008;15: 152-167.

5. Amsterdam JD, McHenry L. The paroxetine 352 bipolar trial: A study in medical ghostwriting. *International Journal of Risk and Safety in Medicine*. 2012;24(4): 221-231.

6. Keller MB, Ryan ND, Strober M, Klein RG, Kutcher SP, Birmaher B, Hagino OR, Koplewicz H, Carlson GA, Clarke GN, Emslie GJ, Feinberg D, Geller G, Kusumakar V, Papatheodorou G, Sack WH, Sweeney M, Wagner KD, Weller EB, Winters NC, Oakes R, McCafferty JP. Efficacy of Paroxetine in the treatment of adolescent major depression: a Randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2001;40:762-772.

7. STI. Paxil-funded publications 1998 to current. 2005. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=znxl0228

8. Wilson D. Drug maker wrote book under 2 doctors' names, documents say. *The New York Times*. 2010 Nov 30: B3

9. Nemeroff CB, Evans DL, Gyulai L, Sachs GS, Bowden CL, Gergel IP, Oakes R, Pitts CD. Double-blind, placebo-controlled comparison of imipramine and paroxetine in the treatment of bipolar depression. *American Journal of Psychiatry*. 2001;158(6): 906-912.

10. Williams, JBW. A structured interview guide for the Hamilton Depression Rating Scale. *Arch Gen Psychiat*. 1988;45: 742-747.

11. Guy W. ECDEU assessment manual for psychopharmacology. US Department of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health, Psychopharmacology Research Branch, Division of Extramural Research Programs. 1976.

12. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Brit J Psychiat*. 1978;133: 429-435.

13. Esfandiari B. Complaint of Scientific Misconduct, Office of Research Integrity of the Department of Health and Human Services. 2011. Available from: http://psychrights.org/Research/Digest/Science4Sale/110708EthicsComplaintAgainstEvansGyualaiNemeroffSachsBowdenetal.pdf

14. Esfandiari B. Complaint of Scientific Misconduct, Office of Research Integrity of the Department of Health and Human Services. 2012. Available from: http://blogs.nature.com/news/files/2012/06/Jay-Amsterdam-today-filed-a-24-page-complaint-with-the-Office-of-Research-Integrity-at-the-US-National-Institutes-of-Health.pdf

15. STI. STI008109, Draft I. 1997 Mar 13. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=frhk0228

16. STI. STI008106, Letter from Grace E. Johnson to Muriel L. Young. 1997 Apr 4. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=frhk0228

17. STI. STI008050, Facsimile from Neil Pitts to Grace Johnson. 1997 May 7. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=grhk0228

18. STI. STI008093, Facsimile from Muriel L. Young to Grace Johnson. 1997, May 19. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=xrhk0228

19. STI. STI008090, Facsimile from Grace Johnson to Muriel Young. 1997, May, 20. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=jrhk0228

20. STI. STI007107**,** Memorandum from Grace Johnson to Marion Philips. 1997, Jun 6. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=krhk0228

21. STI. STI007238, Letter from Sally K. Laden to Charles B. Nemeroff. 1997, Dec 15. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=mrhk0228

22. STI. STI008363, Proposal for a Journal Article by Sally K. Laden and John A Romankiewicz. 1998 Apr 3. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=nrhk0228

23. STI. STI008194, Letter from Sally K. Laden to Cornelius D. Pitts. 1999, Feb 24. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=yrhk0228

24. STI. Email from Sally Laden to Charlie Nemeroff. 1999, Mar 5. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=rrhk0228

25. STI. STI008029, Email from Sally Laden to Neil Pitts. 1999, Jul 14. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=srhk0228

26. STI. STI007834, Email from Sally Laden to Charlie Nemeroff. 1999, Sep 13. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=zrhk0228

27. STI. STI008748, Email from Sally Laden to Cornelius D. Pitts. 2000, Mar 15. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=fshk0228

28. STI. STI008839, Email from Sally Laden to Charles Nemeroff. 2000, Mar 27. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=gshk0228

29. STI. STI008941, Email from Dwight L. Evans to Sally Laden. 2000, May 11. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=hshk0228

30. STI. STI007733, Email Sally Laden to Charles Nemeroff. 2000, Jun 12. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=jshk0228

31. Koerner B. Disorders made to order. *Mother Jones*. 2002 Jul/Aug. Available from: http://motherjones.com/politics/2002/07/disorders-made-order

32. STI. STI007102, Email Sally Laden to Marion Philip. 2000, Jun 13. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=gthk0228

33. STI. STI008644, Email Jack Matthew Gorman to Charles Nemeroff. 2000, Jun 12. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=fthk0228

34. STI. STI008633, Facsimile from Sally K. Laden to Rai Kumar. 2000, Jul 5. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=hthk0228

35. STI. STI008623, Facsimile from Sally K. Laden to Rai Kumar. 2000, Jul 26. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=jthk0228

36. STI. STI007369, Email from Sally Laden to Charlie Nemeroff. 2000, Oct 31. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=nshk0228

37. STI. STI008496, Letter from Charles B. Nemeroff to Jack M. Gorman. 2000, Nov 30. Available from: https://www.industrydocuments.ucsf.edu/drug/docs/#id=pshk0228

38. American Journal of Psychiatry. Information for Authors. *American Journal of Psychiatry*. 2000;157(5): A75-78.

39. Brian D. Project on Government Oversight, Letter to President Obama. 2000. Available from: https://www.pogo.org/letter/2011/07/pogos-letter-to-president-obama-on-bioethics-commission-chairs-ghostwriting-practices/